Dear Dr. Peters,

Thank you for your careful and critical review. While we agree that this paper is a bit of a synthesis, it attempts to quantify known errors and uncertainties in the global C budget while providing a framework for incorporating unknown errors that may be identified down the road. Although this paper may not provide the most sensational results, an appraisal of errors within any scientific discipline is always necessary, especially in the study of the global C budget, where errors are often not reported or are reported in an unsystematic manner.

Please find below our responses to your specific questions in italics.

1. Abstract, line 3. In the first instance write out carbon (C)

*This has been changed*

2. The abstract has a feel that fossil fuel emission uncertainty has "come to dominate", but this seems to contradict Figure 11? It seems LUC still dominates, but FF will dominate soon?

*The abstract has been revised to first focus on how the errors have changed in the various terms in the carbon budget and then how this affects uptake uncertainty. One critical point that we would like to make is that the errors associated with fossil fuel emissions are greater than the total emissions from land-use. We think that the re-worked abstract makes this point more clear.*

3. Page 14934, line 19. What about process emissions (other than cement)

*We do consider other processes in fossil fuel emissions, such as gas flaring, bunker fuels, and international transport. P14937 L2*

4. Section 1.2 discusses atmospheric and ocean, and a paragraph for each. Wouldn't it make sense to split to a section on atmospheric and a section on ocean?

*Good point this has been changed*

5. A sentence which is mentioned a few times "Because fossil fuel emissions are often estimated from energy consumption or production statistics, they are a fairly well constrained economic variable". I don't understand this. Are FF an economic variable? What is a constrained economic variable? And why is something estimated from production statistics well constrained (is that a casual statement, is there a reference?). I think the energy statistics have quite some uncertainty, and may be less bound then differences in emission factors (or perhaps even energy contents in some cases/countries). I think this statement needs to be reconsidered (also in other places in the paper).

*This sentence has been changed to read:*

*'Because fossil fuel emission estimates are derived from economically-constrained energy consumption statistics, the relative errors in fossil fuel emission estimates are fairly small and thought to be between 5 and 10% (Andres et al., 2014). However, because fossil fuel emissions currently account for > 90% of*

*total emissions, even relatively small errors can result in potentially large uncertainties in absolute C uptake calculated at the global scale (Francey et al., 2013)' L258 in revised text.*

*Essentially we are saying that of all the terms in the global C budget fossil fuel emission estimate errors are relatively small because they are estimated from energy statistics which are a variable of economic concern and often related to a nation's gross domestic productivity. However, fossil fuel emissions are the largest emission flux into the atmosphere, so these absolute error numbers are considerable- in this case a small percentage of a big number (e.g. fossil fuel absolute errors) is still bigger than a large percent of a small number (e.g. land use absolute errors)!*

6. Section 2.2.1. Since this is talking about atmospheric concentrations, it would be useful to give numbers here in both ppm and PgC.

*For the sake of consistency we decided to use the same currency of PgC yr$^{-1}$ for all the terms in the global carbon budget. However, we do offer the conversion to allow the reader to go from ppm to PgC . L223 of revised text:*

*'For direct comparison with other terms in the global C budget, molar mixing ratios of atmospheric $CO_2$ are converted to a mass of petagrams (Pg= $10^{15}$g) C using the conversion factor 2.124 PgC ppm$^{-1}$'*

*It would become too confusing if the units were switched for each component of the global C cycle and thus each section of the paper.*

7. Section 2.2: "Because fossil fuel emission estimates are derived from economically constrained energy consumption statistics, errors in these emission estimates are relatively small". As before, how is this economically constrained and how big is small (5%, 10%, 20%)? This also seems to contradict other parts of the text saying that emission uncertainty now dominates.

*See response to comment 5 above.*

8. Page 14941, line 10: Ok to reference Francey et al, but it may be worth also referencing the comment and response to that paper.

*The comment to the Francey paper has been added.*

9. Section 2.2.1. The word "error" is used here a lot. Some of the uses are not really "error"? As an example ("accounting practices") if one country uses a sector approach and another reference approach, is one of them in "error", which this is just a different method to estimate emissions? If cement production is not included then I would only call it an "error" if they wanted to include it, but didn't. Really, not including it is a system boundary question and hence a structural uncertainty?

*Although we have tried to use the term 'error' strictly in a statistical sense to describe estimate errors (ie. ε) and distinguish them from our calculated 'uncertainty' in uptake, we have probably misused the term 'error' in practical speak. This section has been revised to reflect how different reporting practices by different countries can lead to uncertainties in global emission inventories. L270to 281*

10. Page 14942, line 1-2: “due to social and political pressures”. I don’t think Guan et al were that strong, but suggested it as a possible reason.

*This has been removed L278*

11. * Page 14942, line 6+: I am not sure I completely followed this. Countries are grouped to regions, and each region has a specific uncertainty. Ok (though, it would be good to give a table of the uncertainties for each region, helps for reproducibility). I didn’t understand the weighting bit. This is since you take random subset of countries from the region in the bootstrap, and then you need to rescale to replicate the regional total? What is the link to the errors of the largest emitters? I see you reference Andres et al, but I think adding an extra sentence of clarification may help [Incidentally, I have read Andres, and I searched for “Monte Carlo”, “bootstrap”, “weight”, and none of these words came up]. On the constant error “factors” are constant over time, is this the relative error?

*We have added a table of country-level uncertainties (from Andres et al. 2014) for the supplementary materials and section 2.2.1 has been revised to clearly explain the bootstrap error estimates and how they were weighted based on emission estimates.*

12. * I am not an expert on bootstrap methods, but perhaps you need to give a few words on why you are using bootstrapping in this case (or paper). One way to generate samples would be to assume that you would have a relative error for each region (say 10%, specifying a standard deviation) and then assume a distribution (say log normal) and apply a random distribution to generate different samples. Are you doing this, and then resampling? I did not really see how you came up with a distribution.

*For this paper, we created distributions by sampling from the country errors based on the weighted probabilities (see text). This was done 1000 times for each region, with the mean error of all countries being taken each time. The 1000 iterations formed the final regional joint distributions. This method resulted in smoothed distributions when the regions contained countries with different error measurements. Since the smoothed distributions were weighted towards the higher emitters, sampling from the distributions ensured that the region-wide errors were more accurate than simply sampling from the errors for countries within the region.*

13. Page 14943, line 1. Ok, I am perhaps a little slow. But what is El Camino? Google came up with some interesting results, so I guess this is not a standard term? Why did you use it?

*We introduce this term to describe our novel approach to error estimation, whereby errors in the current year are not independent from errors in previous years, thus the temporally correlated errors follow a ‘path’ or ‘camino’. We use this term to distinguish our approach from a conventional monte carlo type approach where the errors are independent in any given year. This has been better explained in the text L 302*

14. Equation 4. I think it is great to include the temporal correlation. But why 0.95? Ok 20 years, but why 20 years as opposed to 10 or 30? I realise there is no data, but some explanation may help. The

correlation will basically give a decaying correlation over time. The correlation with the adjacent year will be 0.95ˆ2, with an inventory 20 years ago 0.95ˆ20=0.35? Is that how I should interpret?

*We acknowledge that the 20 years of autocorrelation is rather arbitrary, but that it is highly unlikely that nations, especially large emitters are going to retroactively correct their emissions after 2 decades and this has been shown in the literature. The main contribution here is the autocorrelation function and not the 20 years. This has been revised to read (L311to 317):*

*'We note that our selection of ~20 years for the persistence of autocorrelation in emission error estimates is somewhat arbitrary; it assumes that errors are not corrected retroactively after 20 years. While it is conceivable that emission errors could be corrected going back even further in time, it has been shown that estimates tend to converge after a decade (Marland et al., 2009) therefore 2 decades is a fairly conservative estimate of the time-dependence of errors. '*

15. * Page 14943, line 9+. Ok to include CDIAC and EDGAR. But why BP. BP has crude estimates with no methodological description. The estimates can sometimes differ substantially at a national level. I would suggest it is better to use IEA, and better still, use IEA sectoral and IEA reference to make a subset of 4 emission estimates. Did you include cement with BP? If not, you will introduce a bias to the results.

*We simply wanted to include 3 independent estimates of fossil fuel emissions, so I think that the BP estimates actually serve as a pretty good independent estimate because they are not estimated by academics but rather from industry, with a whole different set of assumptions and biases. Many of the academic estimates have similar assumptions and conversion factors and accounting practices, so they are not necessarily 'independent'. In fact, while the BP estimates appear to be biased high since 1990, they were biased low during the 70s and 80s. This is perhaps indicative of another important point from this analysis- that the emission errors are time dependent on decadal timescales. It remains to be seen whether BP will adjust their estimate so that they correspond better with CDIAC and EDGAR. We could replace the BP estimates with the IEA estimates (and probably will for future analyses), but replacing these estimates will not change the fundamental conclusion of our paper that fossil fuel emission errors now dominate global C uptake uncertainty. It does not matter if we are considering 5% or 7% of a very large emission estimate the resulting number is still the largest error term in the budget. All of our fossil fuel emission estimates including the BP estimates take into account emissions from cement production. This has been mentioned in the revised methods (L320).*

16. * Equation 6. I will echo my point equation 4, but why 0.05. That is a tiny correlation. It is basically no correlated, and that correlation diminishes over time. Surely the correlation should be larger, even 0.95 as for FF. And how does 0.05 translate to 5 years? From the Global Carbon Project work a change in method can result in a complete change in the time series from 1959. I would expect the uncertainty in LUC to persist much longer than 5 years and certainly no less than the FF.

*Once again the value of persistence is arbitrary here and it is rather the approach that is important. We selected this value based on the benchmark estimates of land use change emissions from Houghton which are updated every 5 years. This has been better explained in the text (L 354to 357).*

17. Page 14946, line 1+. The AF is introduced here, and mentioned a few times throughout. But, there seems to be no reference to the detailed analysis of AF in the literature. In the last 5 years so, several papers have been discussed on this topic, and I think it is worth linking to that literature here.

*Good point! We overlooked that we presented this result and failed to discuss it in the context of the literature. We have added an entire paragraph on AF to the revised discussion (L 656 to 673).*

18. Equation 9. My first reaction was that this was a correlation matrix (use of Sigma), but this just represents combinations of different datasets? (3 FF and 3 LUC leads to 9 combinations?) For each cell in the matrix you have 500 samples (it is like a 3D matrix) and you have 52 years? I guess I am repeating what you are writing, but this suggests the explanation needs a slight tweak. . .

*Equation 9 has been clarified based on these comments and the comments from Reviewer 2.*

19. Page 14949, line 5+. "difficult to determine dC/dt was in fact increasing". This is a little confusing, and I think a bit of care is needed. It is not that you have written anything C7218wrong, but you are talking about the rate of change of a rate of change (dC/dt). C is clearly increasing (dC/dt is positive), but it is unclear if dC/dt is increasing (d2C/dt2). In other words, it is unclear whether the growth in C is accelerating over time? I would just be a little more explicit on some of these distinctions.

*Good point, it is always tricky discussing the derivative of a derivative. This discussion has been simplified and hopefully clarified (L 436 to 443).*

20. Section 3.2. There is again the term "error" used here, and am not sure it is correct. Is "uncertainty" better?

*We think that this is in fact the appropriate term because it reflects the increasing contribution of fossil fuel emissions from developing countries which have a higher relative error as well as an apparent divergence in the individual emission inventories. Striclty from a statistical perspective this represents an increase in the error of the estimate in question (i.e.* $\varepsilon_F$ *). I suppose that we could call it 'decreased precision' instead but this is largely semantics.*

21. Page 14951, line 24+. There is improved detection of changes in C update, but a recent change in that trend. Is this just a trade-off between the constantly reducing uncertainty in dC/dt but the growing uncertainty in E? This "trough" in the last decade may be more a coincidently combination of the uncertainties, rather than anything more physical in the climate system?

*Your assessment of the competing effects of decreased error in dC/dt and increasing errors in E is correct. The text has been modified to reflect this (L 504 to 506).*

22. Page 14952, line 24. Ok, 122 simulations had a decreasing trend in N? That would mean that atmospheric growth (dC/dt) grew faster than emissions? This sounds unphysical, or I misinterpreted. It would be quite interesting to see a plot of the 122 sets of emissions and dC/dt to see if they look physical in any way!

*This statement in the text only applies to net ocean uptake ($N_O$) and suggests that there is a 3% chance that net ocean uptake has not increased. This probably indicates that we have added to much error to the uptake estimates from ocean biogeochemical models, rather than some physical impossibility.*

23. * Discussion. It is ok to have a discussion, but I must admit I had a feeling of deja vu. I think I read some of this before! Perhaps one weakness of the paper is that it does not link to the existing literature. The Global Carbon Project also does quite some work on understanding the global carbon cycle, yet this work is barely mentioned (only mention is to the ocean data?). I think the discussion would be a good place to compare with the work of the GCP. What new is added with your analysis? E.g., "others have underestimated X", "we find that there has been insufficient emphasis on Y", etc. That would greatly improve the discussion

*The discussion has been revised extensively, including an additional paragraph on the airborne fraction and the inclusion of references that place our results in a broader research context.*

24. "The greatest source of error in fossil fuel emission estimates is derived from national energy consumption statistics that can be as high as 20% of total emissions for some nations". But earlier this was not uncertain as it was economically constrained?

*On a relative scale these errors are still much smaller than errors in land use emissions which are on the order of 50% because more people care about fossil fuel consumption than land consumption.*

25. Figure 3. There is a missing something "All inventories also include cement production as"? What did you do for BP?

*This has been revised in figure caption 3 all inventories included cement production.*

26. * Figure 4-6. The figures are generally nice, but these ones make it difficult to get an idea of the distribution. For example, in Figure 4 it looks like a value between 0 and 2.5 is equally likely. Is it possible to plot with shading to give some idea of the distributions? Where is the median? Where are the 1 sigma values, 2 sigma, etc. Alternatively, a set of histograms could be placed under Figures 4-6 to show the distributions.

*We decided to show all of the simulations, instead of obscuring the data by showing the statistics. It is informative for the reader to realize that while it is not likely (in a probabilistic sense) that land-use emissions were negative, which would actually indicate a net uptake of C, based on our simulations it is possible to get negative values. We do show levels of uncertainty once we arrive at our C uptake estimates; however, it is more revealing to show all of the simulations and let the reader decide which simulations are more likely.*

27. * Figure 4-6 (4,6 in particular). This figures show large "spikes" every year. This I imagine is a lack of temporal correlation. If you put in a strong temporal correlation (0.95) then those spikes will disappear. This means that if I plotted an individual realisation in these figures, they would be rather random (the emissions in year t+1 will have no link to the emission in year t). This effect should be much smaller in the fossil emissions. I think it is worth exploring individual realisations a little to see if they make sense.

Ultimately, I would consider increasing correlations in the LUC data (as mentioned earlier). One would also expect correlations in the ocean data. Each measurement or model run is not independent of the previous value, in which case I would expect some temporal structure in that data.

*This is true because we have plotted the simulations as lines the degree of apparent 'spikeyness' is in fact a function of the temporal correlation of errors in the estimates. For instance, the fossil fuel emission estimates appear the least spikey because we have arbitrarily assigned a 20 year autocorrelation function based on observations in the literature (see Marland et al.) compared to the land-use emission errors which only have a 5 year autocorrelation based on forestry statistics that are updated and released every 5 years (see Friedlingstein et al.). However, for the ocean uptake estimates we assigned errors independently for every year because we have no idea how often these models are revised. One could include time dependent errors in the ocean C uptake, but it would entail redoing our entire analysis and while it would result in much smoother error distributions it probably would not change our results substantively because there is very little inter-annual variability in the ocean C uptake estimates to begin with.*

28. Figure 8B, why is it so skewed?

*This is a good question and I am not certain. However, it could be due to the change in variability in global C uptake that is enhanced when we remove the land use emissions that show very little trend over the last 50 years.*

29. Figure 8C, D. It would be good to show the 0 value on these figures.

*Not all of the figures have zeros on the axes, so this is not possible. This is why we color coded the bars , such that negative values indicating increased C uptake from the atmosphere are filled grey.*

30. Figure 9. I like this, it would be good to have colours that contrast more than blue and green (though I see why you chose those colours).

*I think that the blue for ocean and green for land are pretty intuitive to the reader.*

31. Figure 11. Nice summary of the paper. These seems to contradict the finding in the abstract? LUC is still the largest source of uncertainty, but FF is growing very fast.

*Thanks! The abstract has been changed to highlight this point.*